CONDITIONAL AND UNCONDITIONAL INFERENCE ABOUT

THE MEAN OF THE UNIFORM DISTRIBUTION WITH KNOWN RANGE

by

Nuan-Hua Duan and David V. Hinkley
Stanford University University of Minnesota

Technical Report No. 279

January, 1977

CONDITIONAL AND UNCONDITIONAL INFERENCE ABOUT THE MEAN OF THE UNIFORM DISTRIBUTION WITH KNOWN RANGE

Nuan-Hua Duan and David Hinkley

Stanford University and University of Minnesota

ABSTRACT

The statistical concepts of significance tests, confidence intervals and conditional inference are illustrated in the analysis of a uniform distribution with known range.

1. INTRODUCTION

In this note we use the simple case of a uniform distribution with known range to illustrate several concepts and methods of statistical theory. The two-dimensional sufficient statistic contains an invariant or ancillary component, namely the sample range, which leads to particularly simple and appealing conditional inference procedures. The simplicity of the model allows easy and informative illustration of significance tests and confidence limits.

Section 2 summarizes results concerned with the sufficient statistics for the model. Both unconditional and conditional significance tests for the location parameter are derived in Section 3. Confidence limits and intervals are described in Section 4.

The model discussed in this note is of some historical interest in that Welch (1939) used it to demonstrate the "inefficiency" of conditional inference; Bartlett (1940) suggested that conditional methods might be efficient if used in a flexible way. Recently Barnard (1976) has derived the "optimal" confidence interval for the uniform distribution; a simple variant of Barnard's derivation is given in Section 4.

2. SUFFICIENT STATISTICS AND THEIR PROPERTIES

Let $X_1, \ldots, X_n$ be i.i.d. with uniform probability distribution on the interval $[\theta, \theta+1]$, where θ is unknown. If $I_{[a,b]}(\cdot)$ is the indicator function for the interval $[a,b]$, then the likelihood function derived from $X_1, \ldots, X_n$ is

$$\text{lik}\,(\theta; X_1, \ldots, X_n) = I_{[\theta,\theta+1]}(X_{(1)})\, I_{[\theta,\, +1]}(X_{(n)}) , \qquad (2.1)$$

where $X_{(1)} = \min\limits_{i \leq j \leq n} X_j$ and $X_{(n)} = \max\limits_{i \leq j \leq n} X_j$. Thus $S = (X_{(1)}, X_{(n)})$ is minimal sufficient for θ. That S is not complete follows from the fact that the distribution of the range $R = X_{(n)} - X_{(1)}$ is independent of θ. In the usual terminology, R is an ancillary statistic, in particular a location invariant statistic.

Equivalent to S is the pair (M,R), where $M = \frac{1}{2}(X_{(1)} + X_{(n)})$ is the mid-range; note that $M-\frac{1}{2}$ is the unbiased maximum likelihood estimate. Direct calculation shows that the joint probability density of S, and hence that of (M,R), is

$$n\ (n-1)\ r^{n-2} \qquad (\theta+\tfrac{1}{2}r \leq m \leq \theta+1-\tfrac{1}{2}r,\ 0 \leq r \leq 1).$$

Therefore the marginal probability density of R is

$$f_R\ (r) = n\ (n-1)\ (1-r)\ r^{n-2} \qquad (0 \leq r \leq 1)\ , \tag{2.2}$$

independent of θ; and the conditional probability density of M given R=r is

$$g_{M|R}\ (m|r;\ ) = (1-r)^{-1} \qquad (\theta+\tfrac{1}{2}r \leq m \leq \theta+1-\tfrac{1}{2}r)\ , \tag{2.3}$$

i.e. uniform on $[\theta+\frac{1}{2}r,\ \theta+1-\frac{1}{2}r]$.

The support of S is a triangle Δ_θ with vertices $A_\theta = (\theta,\theta)$, $B_\theta = (\theta,\theta+1)$ and $C_\theta = (\theta+1,\ \theta+1)$, as shown in Figure 2.1, where the sets M=m and R=r are also illustrated. The diagonal $A_\theta C_\theta$ generates the doubly-infinite line L as θ varies; the support triangle Δ_θ slides up and down L as θ increases and decreases.

Notice from (2.2) and (2.3) that as n increases S is attracted toward B_θ, but that the conditional distribution of M given R=r does not vary with n.

3. SIGNIFICANCE TESTS

We now discuss some "optimal" tests concerning θ, in each case using the geometry of the support triangles $\{\Delta_\theta\}$ to clarify the forms and properties of the test.

Unconditional Tests

First, consider one-sided tests of the null hypothesis H_0: $\theta = \theta_0$ with alternatives H_+: $\theta > \theta_0$. For any specific alternative θ_A the most powerful size α test rejects H_0 for large values of the likelihood ratio

$$lr_{A,0}\ (s) = lik\ (\theta_A;\ s)\ /\ lik\ (\theta_0;\ s)\quad .$$

But by (2.1), the possible values of $lr_{A,0}$ (s) are 0, 1, ∞ corresponding to the events $\theta_0 \leq y_{(1)} < \theta_A$, $\theta_A \leq y_{(1)} \leq y_{(n)} \leq \theta_0 + 1$, $\theta_0 + 1 < y_{(n)} \leq \theta_A + 1$. Because the last event has zero probability under H_0, its probability when

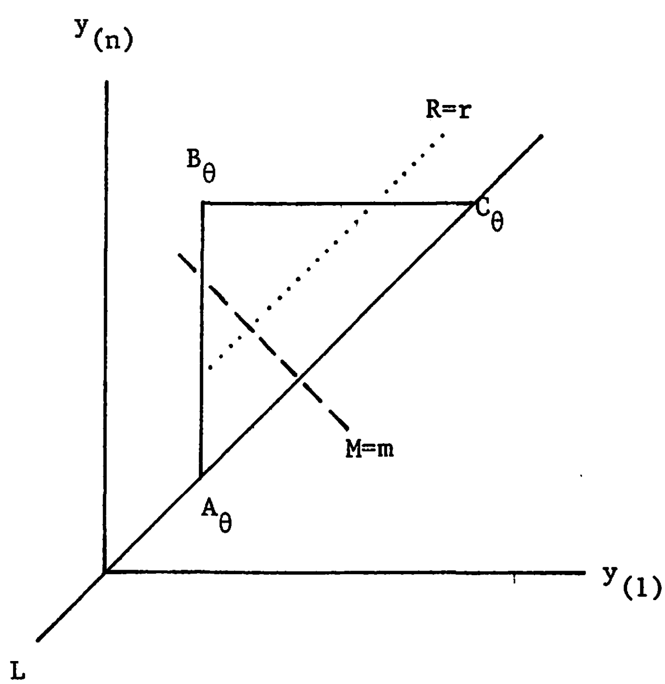

FIGURE 2.1
Support triangle Δ_θ for S

$\theta = \theta_A$ will be a free contribution to power of any test. The size α likelihood ratio test is uniquely defined only if the event $lr_{A,0}$ (s) = 1 has probability equal to α under H_0; the only exactly achievable significance level is

$$\text{pr}\ \{\theta_A \leq Y_{(1)} \leq Y_{(n)} \leq \theta_0 + 1;\ \theta_0\} = \text{pr}\ \{S \in \Delta_{\theta_0} \cap \Delta_{\theta_A};\ \theta_0\}.$$

Therefore, in general, some randomization is required to determine a most powerful exact size α test. This randomization may involve the value of s, and it turns out that such a randomized procedure gives a uniformly most powerful test.

To proceed, we note that any size α likelihood ratio critical region w_α^+ contains a region v_α^+ of size α inside Δ_{θ_0}. For a particular alternative θ_A, the contribution to power from v_α^+ is then

$$\text{pr}\ \{S \in v_\alpha^+ \cap \Delta_{\theta_A};\ \theta_0\} \tag{3.1}$$

because $lr_{A,0}$ (s) = 1 for $s \in v_\alpha^+ \cap \Delta_{\theta_A}$. The total power of the critical region w_α^+ is equal to (3.1) plus

$$\text{pr}\ \{lr_{A,0}\ (S) = \infty;\ \theta_A\}\ .$$

Therefore we can obtain uniformly maximum power by maximizing (3.1) for every $\theta_A > \theta_0$; notice that (3.1) is bounded above by α. The solution for v_α^+ is that size α region contained in $\Delta_{\theta_0} \cap \Delta_{\theta_A}$ for the largest possible value of θ_A. Thus, if we slide Δ_θ up L from Δ_{θ_0} until at $\theta = \theta_1$

$$\text{pr}\ \{S \in \Delta_{\theta_0} \cap \Delta_{\theta_1};\ \theta_0\} = \alpha\ , \tag{3.2}$$

then $v_\alpha^+ = \Delta_{\theta_0} \cap \Delta_{\theta_1}$ makes (3.1) equal to its upper bound α for $\theta_A \leq \theta_1$. But for larger values of θ_A the total power of the resulting test is equal to the maximum value (one), because then the whole of Δ_{θ_A} is in the critical region

$$w_\alpha^+ = v_\alpha^+ \cup \{S: \; y_{(n)} > \theta_0 + 1\} \quad .$$

The optimal region v_α^+ defined by (3.2) is shown in Figure 3.1.

Simple calculation shows that $\theta_1 = \theta_0 + 1 - \alpha^{1/n}$; for example, with $\alpha = .05$, $\theta_1 - \theta_0 = 0.776$ and 0.259 at $n = 2$ and $n = 10$ respectively. Thus the triangle v_α^+ is typically quite large.

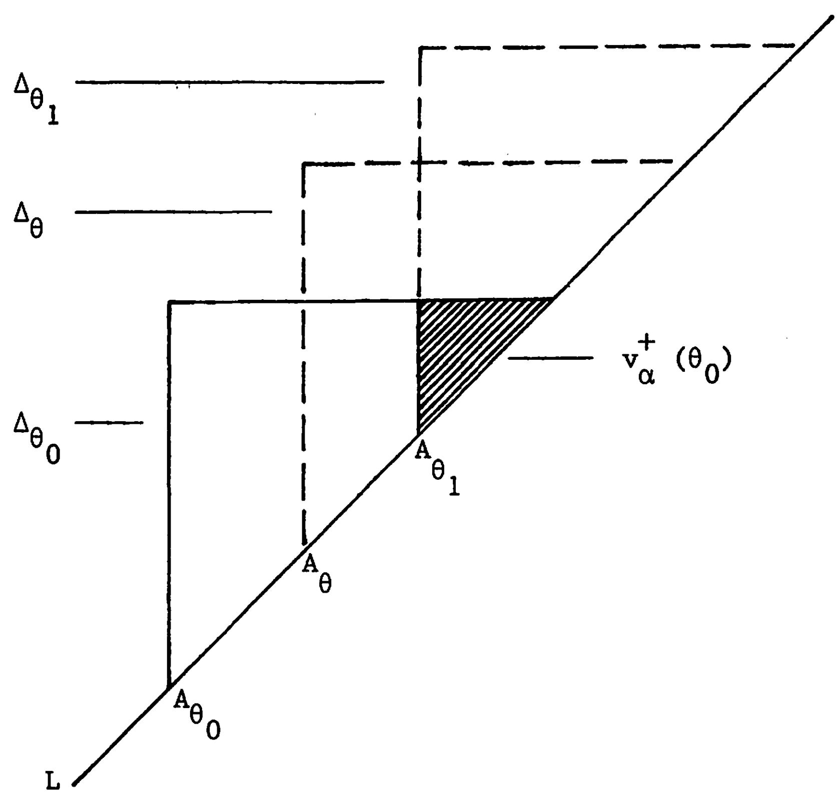

FIGURE 3.1

Test of H_0: $\theta = \theta_0$ versus H_+: $\theta > \theta_0$
Shaded area is intersection of uniformly most powerful size α critical region with Δ_{θ_0}.

A parallel calculation gives the uniformly most powerful one-sided test of H_0: $\theta = \theta_0$ versus H_-: $\theta < \theta_0$, with size α critical region w_α^- intersecting Δ_{θ_0} in a triangle v_α^- that is the mirror image of v_α^+ with lower vertex at A_{θ_0}.

It is not true that the symmetric, unbiased union of one-sided tests, with critical regions $w_{\frac{1}{2}\alpha}^-$ and $w_{\frac{1}{2}\alpha}^+$, is a uniformly most powerful unbiased test of H_0: $\theta = \theta_0$ versus H: $\theta \neq \theta_0$. In fact no such test exists, as we can see readily from the construction of a locally most powerful unbiased test.

The situation is illustrated in Figure 3.2, where the region v_α^* (a translation of the previous v_α^+) sits in the center of $A_{\theta_0} C_{\theta_0}$. The union of v_α^* with the exterior of Δ_{θ_0} is the critical region of the locally most powerful unbiased size α test. That this is so follows from an argument similar to that used above in finding the optimal v_α^+. No region other than v_α^* is both unbiased

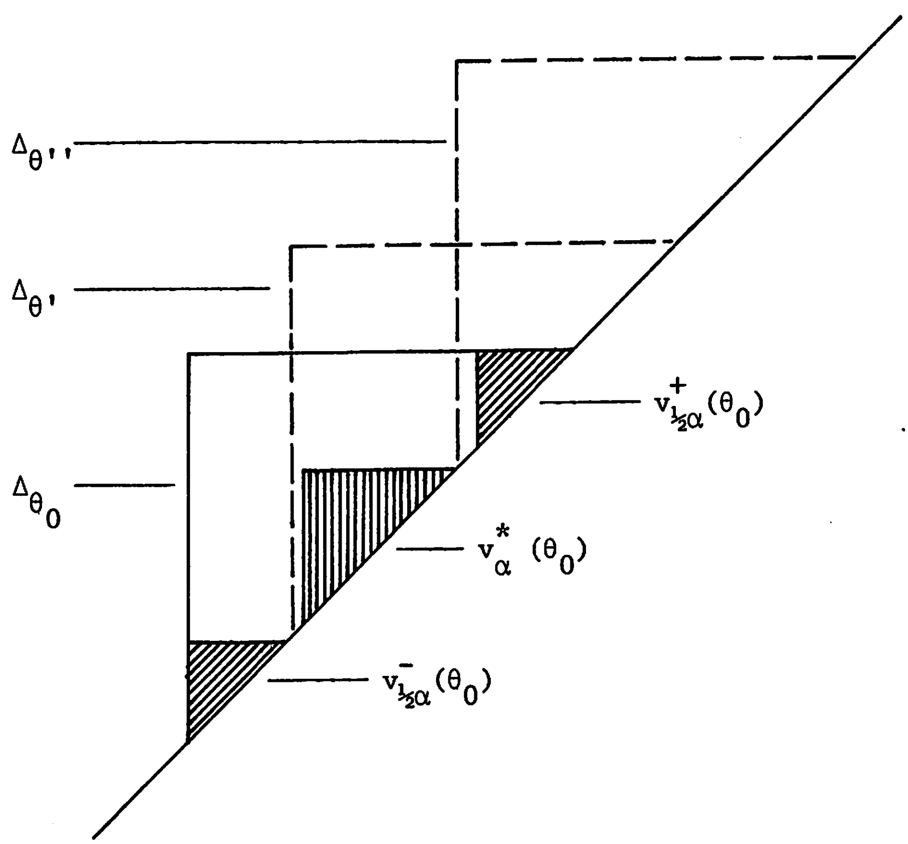

FIGURE 3.2

Unbiased two-sided tests of H_0: $\theta = \theta_0$ versus H: $\theta \neq \theta_0$

and as powerful for such a large set surrounding $\theta = \theta_0$. For the particular value $\theta = \theta'$ indicated in Figure 3.2, the power of v^*_α exceeds that of $v^-_{\frac{1}{2}\alpha} \cup v^+_{\frac{1}{2}\alpha}$ by $\frac{1}{2}\alpha$. However, at $\theta = \theta''$ the situation is reversed. Incidentally one can see directly from the figure that both tests are uniformly unbiased.

In view of the preceding discussion it is worth emphasizing that randomization has been introduced in order to define a test with exact size α. On intuitive grounds one would say that v^*_α is a silly rejection region, because it corresponds to small values of $\hat{\theta} - \theta_0$, where $\hat{\theta} = M - \frac{1}{2}$ is the "best" estimate of θ. Both optimal one-sided tests are intuitively appealing in the sense that large values of $|\hat{\theta} - \theta_0|$ lead to rejection of θ_0. Intuition and the likelihood disagree.

Conditional tests

The tests described so far have a curious property, namely that if R is sufficiently close to 1, then no value of M will lead to rejection of H_0. Thus in Figure 3.1, if the line $R = r$ does not pass through v^+_α, then the conditional size of the test is zero. Not to have rejected H_0 at level α means very little for such a value of r. In the extreme case $r = 1$ we know θ with certainty ($\theta = m - \frac{1}{2}$), so that a size zero test exists with power one! In effect the value of the ancillary statistic R tells us how precisely we can locate θ: the conditional standard error of $\hat{\theta} = M - \frac{1}{2}$ given $R = r$ is proportional to $1 - r$. A canon of Fisherian inference is that all inferences about θ should be made using sampling distributions conditional on the value of such an ancillary statistic.

For the present model, conditional size α tests are based on the conditional distribution (2.3). Thus the uniformly most powerful conditional size α test of H_0: $\theta = \theta_0$ versus H_+: $\theta > \theta_0$ has critical region which overlaps Δ_{θ_0} in

$$u_\alpha^+ = \{s: \ m \geq \theta_0 + \tfrac{1}{2}r + (1 - \alpha)(1 - r)\} \quad . \tag{3.3}$$

Again this is a randomized likelihood ratio test (because likelihood ratio values $lr_{A,0}(s)$ are unaffected by conditioning).

As with the unconditional tests, there is no uniformly most powerful unbiased test of H_0: $\theta = \theta_0$ versus H: $\theta \neq \theta_0$. Again there is a locally most powerful unbiased test, which here has critical region

$$u_\alpha^* = \{s: \ |m - \tfrac{1}{2} - \theta_0| \leq \tfrac{1}{2}\alpha(1 - r)\} \quad .$$

As before, this is counterintuitive.

An obvious criticism of the best conditional size α test is that its unconditional power is by definition less than that of the best unconditional test; Welch (1939) originally voiced this criticism. But there is nothing in the general theory of tests that requires α to be fixed in advance, nor that requires interpretation of significance to be independent of r. Thus in the present case the Fisherian objection to the unconditional size α test would be that <u>after</u> r is observed, the observed significance of m conditional on R = r should be reported, <u>not</u> α. In fact the most informative summary in testing H_0: $\theta = \theta_0$ versus H_+: $\theta > \theta_0$ would be that the data are significant at the level

$$\frac{\theta_0 + 1 - \tfrac{1}{2}r - m}{1 - r} \quad .$$

This reflects the conditional probability of error that would apply if data as extreme as those observed were regarded as grounds for rejecting H_0.

4. CONFIDENCE LIMITS AND INTERVALS

As explained by Cox and Hinkley (1974, Chapter 7), for example, confidence limits and intervals for θ may be obtained simply by inverting significance tests. Thus, for example, a physically natural lower confidence limit for θ with confidence coefficient $1 - \alpha$ is obtained as the largest value of θ_0 not rejected in a size α test of H_0: $\theta = \theta_0$ versus H_+: $\theta > \theta_0$. In terms of the geometry of S, we plot the observed value s and slide Δ_{θ_0} up L until s touches the boundary of the critical region. The unconditional lower limit is then seen to be

$$\max (X_{(1)} - c, X_{(n)} - 1) \tag{4.1}$$

with confidence coefficient $1 - \alpha = 1 - (1 - c)^n$; i.e. $c = 1 - \alpha^{1/n}$.

The curious conditional properties of unconditional tests (Section 3) are reflected here by the fact that if $r > 1 - c$, then (4.1) is a <u>certain</u> lower limit for θ; this was pointed out by Pierce (1973). Conditional confidence limits are obtained from the conditional distribution of M given $R = r$. Thus, for example, a natural confidence interval for θ with confidence coefficient $1 - \alpha$ is

$$[M - \tfrac{1}{2} - \tfrac{1}{2}(1 - \alpha)(1 - R),\ M - \tfrac{1}{2} + \tfrac{1}{2}(1 - \alpha)(1 - R)] \quad . \tag{4.2}$$

Actually, from a strictly inferential, i.e. non-decision, viewpoint the interval (4.2) is "fair" only if $\alpha = 0$ or 1, because the likelihood function is constant over the whole conditional range of θ.

The above confidence limits are obtained via randomized likelihood ratio tests. Suppose that we wished to obtain directly a best confidence interval in the sense of shortest average length. Such an interval would be obtained by inverting a uniformly most powerful test, but no such test exists here. The following derivation of a best confidence interval is similar to that given previously by Barnard (1976).

We require an interval which covers θ with probability $1 - \alpha$ and which has smallest expected length. Let the coverage probability conditional on $R = r$ be $1 - \alpha(r)$. But conditional on $R = r$, <u>any</u> confidence interval with coefficient $1 - \alpha(r)$ has length $(1 - r)\{1 - \alpha(r)\}$ because of the uniform distribution (2.3) of M given $R = r$. Therefore the unconditional expected length of the confidence interval is

$$\int_0^1 \{1 - \alpha(r)\}\,(1 - r)\,f_R(r)\,dr \tag{4.3}$$

and by definition

$$\alpha = \int_0^1 \alpha(r)\,f_R(r)\,dr \quad . \tag{4.4}$$

Our problem is, then, to choose $\alpha(\cdot)$ so as to maximize

$$\int_0^1 \alpha(r)(1 - r)\,f_R(r)\,dr$$

subject to

$$\int_0^1 \alpha(r)\,f_R(r)\,dr = \alpha \quad .$$

Because $1 - r$ is decreasing, we must make $\alpha(r)\,f_R(r)$ as large as possible for small r. Therefore the optimal choice is

$$\alpha(r) = \begin{cases} 1 & (0 \le r \le r^*) \\ 0 & (r^* \le r \le 1) \end{cases}$$

where

$$\alpha = \int_0^{r^*} f_R(r)\,dr \quad .$$

That is, the best confidence interval for θ is

$$[M - \tfrac{1}{2} - \tfrac{1}{2}(1 - r),\ M - \tfrac{1}{2} + \tfrac{1}{2}(1 - r)]$$

if $r > r^*$. Otherwise the confidence interval has zero length, e.g. including the single point $m - \tfrac{1}{2}$. This situation is rather extreme, but notice that it is coherent with the constant likelihood. Of course it would seem wise to give the conditional probability of coverage (0 or 1) appropriate for the observed data, not $1 - \alpha$!

BIBLIOGRAPHY

BARNARD, G.A. (1976). "Conditional Inference is not Inefficient". Scand. J. Statist. 3, 132-4.

BARTLETT, M.S. (1940). "A Note on the Interpretation of Quasi-Sufficiency". Biometrika, 31, 391-2.

COX, D.R. and HINKLEY, D.V. (1974). Theoretical Statistics. Chapman and Hall, London.

PIERCE, D.A. (1973). "On Some Difficulties in a Frequency Theory of Inference". Ann. Statist. 1, 241-50.

WELCH, B.L. (1939). "On Confidence Limits and Sufficiency, with Particular Reference to Parameters of Location". Ann. Math. Stat. 10, 58-69.